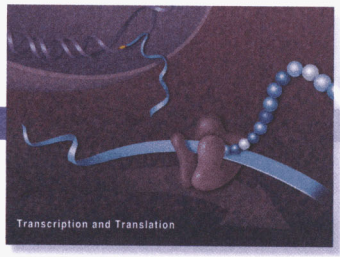
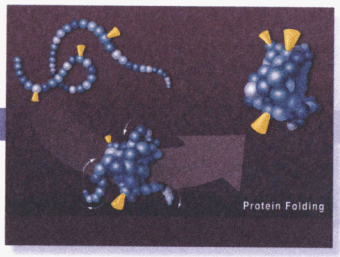
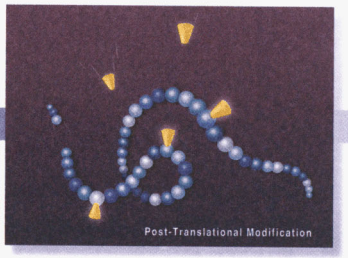
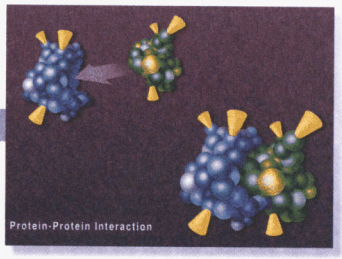
The Promise of Proteomics

Leading the Way to 21st Century Medicine









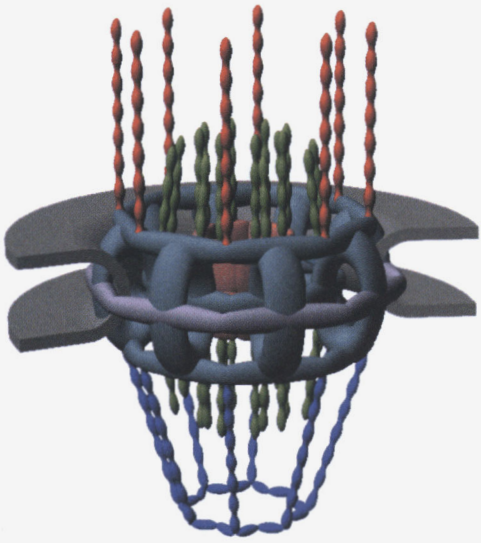
New York Academy of Sciences

This report was prepared for the New York Academy of Sciences by Vicki Brower. It summarizes presentations and discussions at a media backgrounder on "The Promise of Proteomics," sponsored by the New York Academy of Sciences and held at NYAS headquarters in New York City on January 31, 2001. The backgrounder was webcast live and is currently archived on the NYAS website, www.nyas.org.

Made possible by an educational grant from the Applera Charitable Foundation

The Applera Charitable Foundation, Inc. was created in 1999 to foster education among professionals and the public in the science of genomics. This focus allows the Foundation to use its resources to sponsor educational programs in genomics, proteomics, and related technologies.

The Promise of Proteomics



Leading the Way to 21st Century Medicine

Above: Nuclear pore complex, courtesy of Drs. Michael Rout and Brian Chait, The Rockefeller University

New York Academy of Sciences

PARTICIPANTS

N. Leigh Anderson, Ph.D.
President and CEO
Large Scale Proteomics Corporation
9620 Medical Center Drive, Suite 201
Rockville, Maryland 20850-3338 USA
T:301.926.8229; F:301.926.4736;
E:Leigh.Anderson@lsbc.com

Brian T. Chait, D.Phil.
Camille and Henry Dreyfus Professor
The Rockefeller University
Laboratory of Mass Spectrometry and
Gaseous Ion Chemistry
1230 York Avene
New York, New York 10021 USA
T:212.327.8849; F:212.327.7547;
E:Chait@mail.rockefeller.edu

David Cowburn, Ph.D., D.Sc.
President & CEO
New York Structural Biology Center
C/O Box 163, The Rockefeller University
1230 York Avenue
New York, New York 10021-6399
T:212.327.8270; F:646.349.2840;
E:cowburn@rockefeller.edu

Barbara J. Culliton
Chairman and President
Applera Charitable Foundation
c/o Celera Genomics
45 West Gude Drive
Rockville, Maryland 20850 USA
T:240.453.3200; F:240.453.3250;
E:Barbara.Culliton@celera.com

Denis F. Hochstrasser, M.D.
Professor, Medical Biochemistry Department
Faculty of Medicine
University of Geneva
Geneva University Hospital
Geneva 14, CH-1211 Switzerland
T:41.22.372.7355 or 56
F:41.22.372.7399;
E:Denis.Hochstrasser@dim.hcuge.ch

Joshua Lederberg, Ph.D.
Sackler Foundation Scholar and
President Emeritus
The Rockefeller University
Laboratory of Molecular Genetics
and Informatics
1230 York Avenue
New York, New York 10021 USA
T:212.327.7809; F:212.327.8651;
E:jsl@mail.rockefeller.edu

S. Leslie Misrock, LL.B.
Senior Partner
Pennie & Edmonds LLP
1155 Avenue of the Americas
New York, New York 10036-2711 USA
T:212.790.6428; F:212.997.7417;
E:misrocks@pennie.com

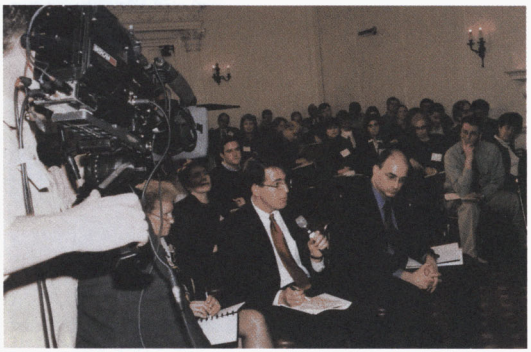
Rodney W. Nichols
President and CEO
New York Academy of Sciences
2 East 63rd Street
New York, New York 10021 USA
T:212.838.0230 ext. 200
F:212.753.3479;
E:rnichols@nyas.org

John H. Richards, Ph.D.
Professor of Organic Chemistry and
Biochemistry
California Institute of Technology
Division of Chemistry, 147-75
Pasadena, California 91125 USA
T:626.395.6040; F:626.568.9430;
E:jhr@cco.caltech.edu

B. Michael Silber, Ph.D.
Director, Pharmacogenomics and Clinical
Biochemical Measurements
Pfizer Global Research and Development
Eastern Point Road
Groton, Connecticut 06340 USA
T:860.441.5750; 860.715.1841
F:860.441.1273;
E:B_Michael_Silber@groton.pfizer.com

THE PROMISE OF PROTEOMICS

Proteomics is a one-word answer to the question, "What comes next now that the human genome is sequenced?" In February 2001, Celera Genomics and the Human Genome Project each published their sequences of the human genome. That monumental accomplishment attracted attention to an equally complex stage of research: proteomics, the study of the proteins coded by genes.



A webcast camera catches the Q&A portion of the proteomics media backgrounder held at the New York Academy of Sciences.

"The proteome refers to the whole body of diverse proteins found within a given organism," said Joshua Lederberg, Nobel laureate and president emeritus of Rockefeller University, at a January 2001 conference at the New York Academy of Sciences. At this conference, sponsored by the Applera Charitable Foundation, experts convened to discuss the promise of proteomics and its challenges in developing new treatments for disease.

Proteomics is the systematic cataloging, separation and study of all the proteins produced by genes within each cell as well as the complex interactions among proteins that ultimately result in health or disease. "Proteomics addresses the question of what proteins do in a cell in a global, integrated way," said Brian T. Chait, professor and head of the mass spectrometry and gaseous ion chemistry laboratory at Rockefeller University.

The terms "proteomics" and "proteome" only came into use between 1995 and 1998 by analogy with genomics and genome, and are still not in standard dictionaries, noted Lederberg. "We know tens of thousands of proteins at this stage, but just as the periodic table enabled us to say that there was a larger number of elements that had yet to be discovered, that's true currently about proteins," he said.

Proteins are encoded in DNA – the genome – which produces RNA, which in turn make proteins (see illustrations on pages 8,9). The proteins are composed of amino acids which are arranged according to particular sequences, which correspond to the sequence of nucleotides in the gene. After synthesis, many proteins are also chemically modified. At this stage, the protein essentially has all the information necessary to adopt its 3-dimensional structure.

Proteins are the workhorse of the cell, and all proteins work together in a complex network to give function, said Denis F. Hochstrasser, professor of medical biochemistry at the faculty of medicine at the University of Geneva, Switzerland. "Genes are the 'blueprints' for information required for life, but proteins expressed in different cells are the dynamic machines responsible for function," added John H. Richards, professor of organic chemistry and biochemistry of the California Institute of Technology.

One of the mysteries of life is how we manage to be so complex with so few genes. With only 30,000 or so human genes estimated by both Celera Genomics and the Human Genome Project – rather than previous estimates of 100,000 or more – it has become apparent that our complexity is tied *not* to the number of genes, but rather to the complexity of their products, the proteins.

Drosophila, the fruit fly, possesses 13,000 genes, and *C. elegans*, the roundworm, has 19,000 genes, compared to humans' 30,000; all three organisms share many homologous genes. The key to each organism's uniqueness lies in the fact that each gene may produce more than one protein – anywhere from six to 20 or more – not the old count of a one-to-one ratio of genes to proteins.

Denis F. Hochstrasser

"The human proteome diversity is tremendous, several orders of magnitude greater than the genome diversity," said Hochstrasser. He estimated that based on 30,000 genes, each expressing 5-6 proteins normally (and up to 20 with increasing age), a human may express up to a half-million proteins. Thus far, scientists have identified only about 80,000, and only a small fraction of them have been studied in detail. This is not surprising in view of the complexity and diversity of proteins. However, the pace of discovery in proteomics has accelerated and the field is entering an exciting new era.

"Clinical proteomics is important in the future development of bio-markers for diagnosis and drug development. Just one type of protein floating in blood can help predict a disease."
--Hochstrasser

How do genes produce such a diversity of proteins? This is partly due to the way instructions for protein syntheses are transmitted from DNA to the cytoplasm. Information for making proteins in the DNA is arranged in blocks – called exons – which are interrupted by other sequences – called introns – whose function is not understood. In making the messenger RNA (mRNA) – which serves as the template for protein synthesis – the cell edits the RNA copy assembled on the DNA by removing sequences that correspond to the introns, a process called RNA splicing. Splicing, it turns out, is more complex than a simple linear editing process; cells can use alternative splicing schemes to generate a variety of messages for a given sequence of DNA. In other words, a linear sequence of exons 1,2,3,4,5,6,7 can not only generate mRNAs whose sequence is 1-2-3-4-6-7, but also other species of mRNAs with sequences such as 1-2-3-6, 1-3-4-5-7, 1-2-4-6, etc. Thus, using alternative splicing, the same DNA sequence – the same gene – can result in a number of proteins.

Protein diversity also arises because proteins are frequently modified after synthesis. For example, two proteins may be covalently linked through their sulfide groups to form a dimer. Or phosphate or sulfate groups may be added. To many proteins destined for the cell surface, a variety of carbohydrates are added. Some proteins are made as an inactive precursor, which is then broken down by enzymatic cleavage to smaller proteins which have physiological function. Thus, by alternative splicing schemes and by a variety of post-translational modifications, the relatively modest number of genes gives rise to a very large number of proteins.

A Wealth of Targets

Speaking at the New York Academy of Sciences conference, N. Leigh Anderson, president and CEO of Large Scale Proteomics Corporation, noted that disease is a malfunction of physiological pathways. "The fundamental underlying cause of disease is that proteins – these nanomachines that do all the jobs in the cell – get out of kilter." Drugs work by correcting protein malfunction – by increasing or decreasing their amounts or by altering their interactions – and thus, by studying proteins, we expect not only to understand the nature of disease, but also learn to design drugs that are more effective than drugs developed using existing approaches.

"A single gene can give rise to a very large number of proteins, and it's the proteins that determine the complexity of organisms."
--Richards

In new drug development there are a number of crucial points in which proteomics can guide scientists. "First would be the identification and selection of good targets compared to superfluous ones," said B. Michael Silber, director of pharmacogenomics and clinical biochemical measurements at Pfizer Inc. A drug target is a protein that might be inhibited or activated to produce a therapeutic effect. "Drugs are only as good as the targets selected," Silber added. "I look at proteomics as a group of technologies that may be able to leverage what we're really trying to accomplish at Pfizer – the discovery and development of new medicines."

N. Leigh Anderson (left) and B. Michael Silber

Speakers debated how many proteins – or, to be more precise, specific sites in proteins – can be targets for therapeutically valuable compounds. The estimates vary widely, from 1,000-5,000 to as many as 10,000-20,000. Currently, the pharmaceutical industry is working with no more than 400-500 targets, many of which are receptors of only one particular type. Thus, the field is potentially wide open. However, the number of potential targets scientists are finding with proteomics vastly outweighs the possibility of investigating every one of them, Lederberg pointed out. "We would be spending many times the GNP if we were to pursue all of them, so there will need to be triage," he added. It will be essential to determine which targets are most urgent, most important, and most accessible to research.

Complicating the picture is the fact that different people express variant versions of the same protein; Cytochrome P-450, for example, is responsible for metabolism of many drugs in the body and thus helps explain why the same drug may affect some patients differently. Variations in the same protein are also known to be involved in individual susceptibility to some diseases. The number of targets, then, is actually amplified by the number of variants of each gene and the proteins it produces. In fact, the number of targets is probably even larger than the total number of proteins produced because proteins interact with each other and form larger protein complexes that can be targeted specifically.

A method scientists are using to home in on targets, as in genomics, is comparing homologous genes and proteins among species to understand function. "If you can find a protein in *Drosophila* that does something, it may do a similar sort of thing in mice, in humans, and so on," Richards said. One example is a gene that affects eye development, called Pax 6. If that gene is compromised in the fly, in the mouse, and in humans, they are all blind. Such a finding could give researchers a strong clue that a certain gene and its proteins might be a worthwhile target to pursue.

> ***"Genes contain the information required for life, but proteins make things happen. Proteomics rounds out genomic information by creating a comprehensive picture of genes' ultimate effects. In essence, it gives us a better understanding of all the intricacies – and all the beauty – of biology."***
>
> *--Richards*

The validation of targets, however, remains exceedingly complex at this stage, said Pfizer's Silber. It is necessary to know much more than sequence information or even the presence of unique protein signatures. What is really necessary is to understand proteins' structure and function, and to examine the interplay of the multiple targets being up and down regulated – a "systems" approach to medicine.

The Need for A New Paradigm and Technologies

> ***"Proteomics is the primary tool for really looking to see what drugs do: manipulate a target."***
>
> *--Anderson*

Proteins are too numerous, diverse, and interactive to be studied by a single technique, the speakers at the Academy conference said. To study proteomics, the approaches and tools developed to discover and mine genomic information are being adapted, and sometimes combined with the older, established methods. Researchers are discovering the limits of available technology and are working to find ways around those by developing new approaches and technologies.

John H. Richards

"One of the challenges facing us once we identify proteins is finding out how they are related, and that's something we need to understand if we're going to understand the role of particular proteins," Richards said. It is necessary to discover which proteins interact with which others, and what pathways they follow. Adding to the complexities of proteomics studies is the fact that proteins are dynamic. Rather than a snapshot, we need a movie to capture their action, said Richards. What is certain is that there must be a tremendous evolution in technology to meet the challenges proteomics presents.

Given the opportunities, many academic researchers and pharmaceutical and biotech companies have come together recently – and others are doing so at a rapid rate – to form collaborations to combine disparate technologies and other capabilities. For example, Anderson's company, Large Scale Proteomics, and its partner Biosite Diagnostics are collaborating to develop protein chip arrays, also called antibody arrays, as tools to measure large numbers of proteins in biological samples. Scientists expect such chips to become the preferred technology for high-volume applications of clinical research, diagnostics, and toxicology. Celera Genomics is comparing proteins expressed – both types and amounts – in healthy and diseased tissue, as well as drug-treated and non-drug-treated samples, to find drug targets for a range of diseases. Celera is also automating protein analytical procedures with a goal of analyzing up to 1 million proteins per day. In another strategic alliance, three companies – Myriad Genetics, Hitachi and Oracle – have joined forces, with an investment of up to half a billion dollars, to identify all human proteins and all their interactions. Incyte Genomics and Genicon Sciences are also working to detect infinitesimal levels of proteins present in tissues using Incyte's antibody array products.

Joshua Lederberg

In addition to such collaborations aimed at technology development, several companies are focusing on ways to organize, manage, and possibly mine the available information. For example, Celera is collaborating with Geron in Menlo Park, CA in its work with stem cells, among other projects. A more extensive effort is LSP's Human Protein Index (HPI) Database, whose aim is to inventory proteins occurring in all major human tissues.

"Our initial analysis of the HPI has turned up new candidate diagnostic markers of tissue damage, as well as fascinating insights into the differences between tissues in different regions of the brain," Anderson said. He estimates that on the basis of results of protein characterization from Large Scale Proteomic's high-throughput mass spectrometry facility, the HPI currently covers protein products of 18,000 human genes in 137 different human tissues. "We are in the process of mining that data to determine what proteins are characteristic of different regions of the brain, heart muscle, the liver and kidney, for example. That's the beginning of the translation of the results of this field into medical benefit," he said. Other databases are expected to be announced within the next year or so with identification and links to genes.

(continued on page 10)

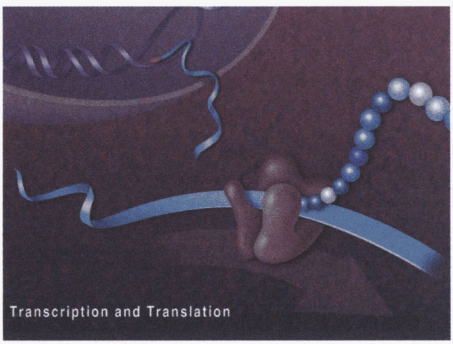
> ***"We know tens of thousands of proteins at this stage, but just as the periodic table enabled us to say that there was a larger number of elements that had yet to be discovered, that's true currently about proteins."***
>
> *--Lederberg*

The Four Steps to Protein Formation

A typical cell may contain thousands of proteins at any time.

Proteins play a variety of roles in the cell. A large class of proteins, called enzymes, plays an essential role in catalyzing all biochemical reactions in the cell – whether they are formation of complex biological molecules or the breakdown of such molecules. Other proteins, such as actin, play a structural role in the cell and give cells shape, help form compartments in which different cellular functions are partitioned, and bind with nucleic acids and other cellular constituents. Some proteins also function as hormones (for example, insulin) and antibodies, and some are involved in transportation of other molecules (for example, hemoglobin, which carries oxygen in red blood cells).

Proteins are formed in the cells in a rather complex multi-step process, which can be divided into four major parts:

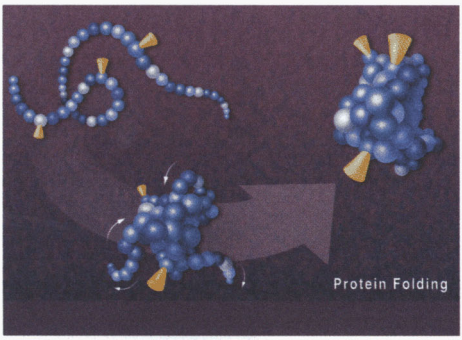


Transcription and Translation

1. Transcription and Translation: Protein Synthesis

All genetic information in the cell is carried in DNA (the genome) which resides in the nucleus. To make proteins, information from the DNA has to be first abstracted. This is done by making a copy of the appropriate parts of the genome by a process called transcription. Transcription and some steps immediately following it result in the formation of a molecule of messenger RNA (mRNA), which has the necessary information to form the protein and serves as the template. Information from the DNA has now been transferred to mRNA.

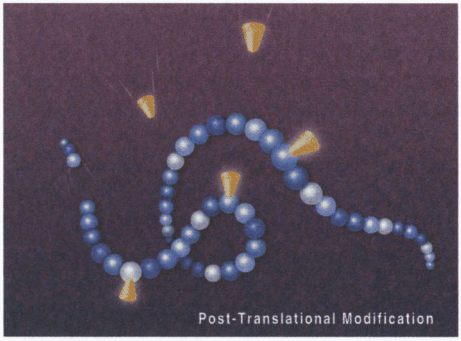
The mRNA leaves the nucleus for the cytoplasm where it binds with a cell component called ribosome – it is here that proteins are assembled in a process termed translation. Once the mRNA is bound to the ribosome, a small variety of cytoplasmic RNA, called transfer RNA, start to bring amino acids one at a time to the ribosome, as though making a necklace with one bead at a time. As more and more amino acids are brought together and linked up, the ribosome glides along the mRNA as if it were a monorail on its track. The sequence of the mRNA dictates which of the many tRNAs bind at any specific point. Thus, the sequence present in the DNA, via the mRNA intermediary, directs the synthesis of protein with a precise and predetermined sequence. Special signals on the mRNA indicate where protein synthesis begins and where it stops.



2. Protein Folding: Generation of Functional and Active Proteins

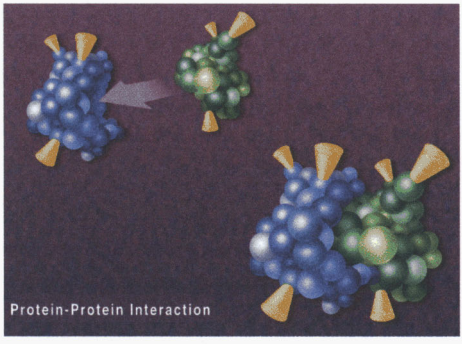
Proteins have distinctive three-dimensional shapes – many are globular, some are fibrous and others have other shapes. The physiological environment plays a role in protein folding. Some proteins fold to expose a hydrophobic (water repelling) region which helps them bind with membranes. Folding and the ultimate three-dimensional structure is facilitated by the formation of covalent bonds (as between two sulfur containing amino acids) and by a variety of non-covaent interactions among the amino acids.

3. Post-Translational Modification: Protein diversification by the addition of sugars, phosphates, and other molecules



Most proteins are modified after they are made by the addition of sugars (glycosylation), phosphate (phosphorylation), sulfate and a few other small molecules. Such modifications often play an important role in modulating the function carried out by the protein. Some proteins are rendered active or inactive by such modifications. Certain proteins – such as membrane proteins – acquire their immunologic properties due to glycosylation.

4. Protein-Protein Interaction: Protein function through binding or formation of complexes



Some proteins work by themselves – they need no company. However, others work only when they are in a complex – bound with other molecules of self, or other proteins or cellular constituents. Hemoglobin, which carries oxygen in the red blood cells, is a complex consisting of four molecules of the protein hemoglobin. Proteins form complexes by binding along surface clefts created by folding in a particular fashion, as well as by ionic and other non-covalent interactions. Protein complexes are a very active area of research.

Using Proteomics for Diagnosis and Prognosis of Disease

"Proteomics is the study of where each protein is located in a cell, when the protein is present and for how long, and with which other proteins it is interacting. It means looking at many events at the same time and connecting them."

--Chait

Proteins are currently used in medicine as biomarkers for diagnosing and staging disease and for determining prognosis. As more efficient, precise techniques are developed to discover and analyze proteins, better methods of diagnosing, treating and preventing illness will follow as the understanding of the molecular basis of disease improves.

"Prognosis is more crucial than diagnosis," Hochstrasser asserted. "Identifying crucial proteins can help to determine a patient's prognosis and select a treatment." One example: Antman and coworkers at Harvard's Brighams and Women's Hospital in Boston reported in the *New England Journal of Medicine* in 1996 (Vol. 335, pp. 1342-9) that by measuring troponin-I, a cardiac-specific blood protein, they could predict a heart attack patient's risk of dying in the next 42 days. "If they found a level below 0.4 nanograms per milliliter in your blood when you come into the emergency room with chest pain, your risk of dying within 42 days is less than 1%. If you have a level greater than 9, the risk is almost 8%. So just measuring one protein in the blood – and the right one – can help a physician determine whether you go home or stay in the ICU to do something about your coronary artery disease."

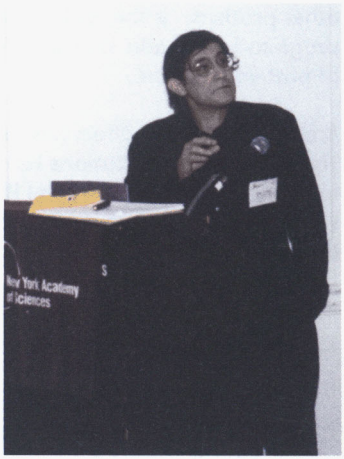
Identifying key proteins in disease can also guide drug development and treatment selection, said Hochstrasser. He described research – also published in the *New England Journal of Medicine* in 1999 (Vol. 340, pp. 1623-9) – conducted by Hamm and colleagues. They studied troponin-T, a close relative of troponin I, which has also been investigated in heart attack patients whose coronary vessel was obstructed. They treated patients with a therapeutic antibody to prevent platelet aggregation and obstruction of blood vessels. They found that antibody treatment had no particular effect in patients with low levels of troponin-T. However, in patients with high levels of troponin-T, who were at high risk for recurring myocardial infarction, the antibody was quite effective in preventing subsequent myocardial events. This study implies that it is appropriate to limit treatment with the antibody only to those patients who have high levels of troponin-T because they are the ones who would benefit from the drug. This ability to match individual patients and specific drug-treatment has a tremendous future potential.

There are other examples where specific proteins or enzymes have proven very useful for treatment or prognostic purposes. However, one of the major obstacles to wider applications of proteomics in clinical medicine is the slow speed of protein separation and identification in the laboratory and the difficulties in dealing with more than very few proteins at a time. "We need faster and better ways to analyze proteins on a large scale," said Hochstrasser. Several research groups at academic and industrial centers are developing innovative methods to overcome these obstacles. Such efforts are aimed at automating the methods of protein separation and analysis with large databases which offer the opportunity to compare the generated data with a large body of existing information for rapid comparison, characterization and possible identification.

As with genomics, this mingling of the "wet" and "dry" laboratory methods is expected to prove extremely useful. Hochstrasser and his colleagues at the University of Geneva, Switzerland, are developing such a system. A powerful molecular scanner, which uses a technique called multi-time of flight, is at the core of the analytical "wet lab." The data generated by the molecular scanner are fed into a protein database for analysis and identification.

As these and other approaches to make proteomics studies faster, better and cheaper continue at a very rapid pace around the world, scientists interested in clinical applications of proteomics look to a future where patient care will be directly linked to specific and individualized information. The ultimate goal is to be able to send a biopsy specimen of a suspicious lesion from an individual patient to the lab – or even better, perhaps a saliva or skin sample – to determine whether it shows cancer, and if so, what type, what stage, and to which drugs it will be sensitive. While some of this can currently be done for certain conditions, the combination of information that will be obtained from genes and proteins is expected someday to make patient care more scientific and therapies more specific and effective as compared to today.

Analyzing Protein Structure and Function



Brian T. Chait

Since each human cell may contain tens of thousands of proteins, scientists must determine where and how much of a certain protein is present and for how long, and with what proteins it is interacting. The structure of individual proteins and the shapes and topology of their interacting complexes are under active investigation. However, to accelerate the pace of such investigations, numerous new tools need to be developed.

A central goal of proteomics is to devise tools that will help scientists analyze cellular function, which they expect will lead to a better picture of normal processes as well as of disease mechanisms. "In a system as complicated as a multicellular organism, you have to be able to look at the entire system in an integrated way," said Rockefeller University's Chait. Proteins operate in complex partnerships with each other and various constituents of the cell, so it is necessary also to track the interactions and changes the proteins produce, rather than viewing individual proteins in isolation.

One method of studying protein interactions is to tag them with a sort of "molecular Velcro" which allows one to pull out the tagged protein together with its strongly associated partners, said Chait. The proteins can then be identified by mass spectrometry. "One must also determine the type of the interacting complexes, the shapes of the individual proteins and modifications that regulate the function of the proteins.

"This is a big job that requires many types of instruments," Chait said. Major efforts are underway to determine the three dimensional structure of proteins using x-ray crystallography and nuclear-magnetic resonance (NMR) spectroscopy.

In a significant move, the National Institute of General Medical Sciences has initiated a new five-year program to fund researchers from a range of institutions working in structural genomics. The centers are located at the Argonne National Laboratory, Lawrence Berkeley National Laboratory, Los Alamos National Laboratory, Rutgers University, Scripps Research Institute and the University of Georgia.

The goal of the NIGMS program is to develop an inventory of all the protein structure families that exist in nature. The centers are seeking to streamline and automate X-ray crystallography and nuclear magnetic resonance spectroscopy. The researchers are beginning by organizing all known proteins into structural, or fold, families based on their genetic sequences. Then they will determine the structure of one or more proteins from each family. Scientists will thereby be able to use gene sequences to approximate structures of all other proteins. If the approach works, this ambitious program would add significantly to our understanding of protein structure.

David Cowburn

Another significant effort has been initiated in New York City. Nine institutions have come together to form the New York Structural Biology Center (NYSBC), which will initially concentrate on using ultra-high field nuclear magnetic resonance spectroscopy to unravel protein structure and function.

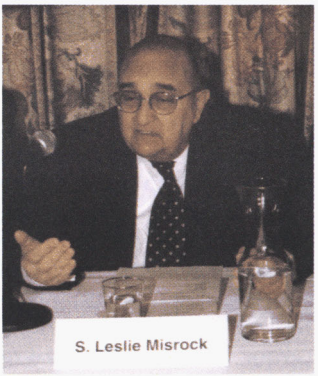
"The NYSBC is unusual because the consortium members have jump-started operations of the Center with their own funds, recognizing the broad opportunities available, and the need to act immediately to retain their competitiveness in this new field," explained David Cowburn, executive director of the Center. The Center's research will focus on membrane proteins, high speed structure determination, flexibility and mobility of multi domain systems, and screening for early lead drug candidates, all of which are key technologies to meet many of the challenges of proteomics, Cowburn said. In Spring 2001, the New York Office of Science, Technology and Academic Research awarded a $15 million grant to NYSBC.

"All proteins work together in a complex network to give function. It's not the protein that functions by itself, but the network of proteins that give the function. There is a feedback loop where the network goes back on protein and the protein on RNA and DNA, and so on. It's the network that gives the complexity."

--Hochstrasser

Despite the many challenges in studying proteins, and the predicted long ramp-up time for the new efforts, a few drugs are already on the market that have the distinction of having been designed on the basis of understanding the structure of the target protein. Agouron and Vertex have developed anti-HIV drugs and Roche has developed an anti-influenza drug. Many other drugs are in the development pipeline.

Other Challenges



S. Leslie Misrock

Many speakers at the New York Academy of Sciences conference spoke about the scientific and technical challenges that need to be overcome to utilize fully the power of proteomics. S. Leslie Misrock, an intellectual property attorney who is a senior partner at Pennie and Edmonds, a New York City-based law firm, shed light on a particularly important issue. He cautioned the audience about the potential for a web of intellectual property disputes that could imperil the biotech and pharmaceutical industry.

He said that there are currently some 7,000 patent applications pending at the United States Patent and Trademark Office; out of these, some 1,000 are subject to contests between various parties to determine invention rights. Citing examples from the 1930s and 1940s in the petroleum industry, he said there are historical guidelines on how to avoid these types of disputes. For example, the oil industry pooled resources and created new institutional arrangements, outside the court system, to help resolve intellectual property disputes. He urged members of the biotechnology and pharmaceutical industry to develop similar mechanisms to curb the emerging tide of intellectual property litigation. If efficient means to resolve such disputes are not found, "the promise of proteomics can turn into a cataclysmic failure," Misrock said.

"If efficient means to resolve (intellectual property) disputes are not considered, the promise of proteomics can turn into a cataclysmic failure."
--Misrock

Other challenges for proteomics center on the limits of current technology in working with membrane and other matrix proteins that are difficult to solubilize. For example, even the answer to a straightforward question, such as the number of proteins in plasma or blood, is not clear. "We don't know whether the total number of proteins is more or less than what we can observe with today's technologies – so, we don't know whether there are proteins present that we can not detect because of limits to the sensitivity of currently available assays, or whether the proteins are just not there?" said Richards.

With the advance of proteomics will come more individualized medicine, predicted Misrock. There is not likely to be a single drug for one disease, but possibly a wide range of treatments for diseases based on their molecular "fingerprint." Proteomics is one area of great promise that researchers will mine for years to come to develop such medicines.

BIBLIOGRAPHY

There have been numerous articles on proteomics in the popular press as well as in scientific newsletters and journals. We provide below a list of some general articles; this is not meant to be a comprehensive or complete list, but only to help in probing this area further.

Anonymous. After the Genome: Proteome. *The Economist Technology Quarterly*. Pp. 27-29. December 9, 2000.

Borman, S. Proteomics: Researchers Use a Variety of Tools to Probe Protein Function and Interactions, with Drug Discovery the Major Goal. *Chemical & Engineering News*, Pp 31-37. July 31, 2000.

Blackstock, W.P., Weir, M.P. Proteomics: Quantitative and Physical Mapping of Cellular Proteins. *TIBTECH*. Vol 17, Pp. 121-127. 1999.

Dreifus, C. A Gene Detective on the Trail of Alzheimer's. *New York Times*. Pp D3. Dec. 5, 2000.

Edwards, A.M., Arrowsmith, C.H., Des Pallieres, B. Proteomics: New Tools for a New Era: Bridging the Gap between Genomics and Drug Discovery. *Modern Drug Discovery*, Pp. 35-44. September 2000.

Elmer-DeWitt, P., Lemonick, M.D., and others. The Future of Drugs. *Time*. Pp. 57-116. January 15, 2001.

Garber, K. Proteomics Gears Up. *Signals*. October 10, 2000. (http://www.signalsmag.com)

Garber, K. The Next Wave of Genomics Business. *Technology Review*. July/August 2000.

James, R. Differentiating Genomics Companies. *Nature Biotechnology*. Vol. 18. Pp. 153-155. 2000.

Kalb, C. A Cancer 'Smart Bomb.' *Newsweek*. Pp. 68. December 18, 2000.

Licking, E. Beyond the Genome: Biotech's New Holy Grail. *Business Week*, April 10, 2000.

Parekh, R. Proteomics and Molecular Medicine. *Nature Biotechnology*. Vol. 17, Pp. 19-20. 1999.

Razvi, E., Latterich, M. Technology Trends for Proteomics and Beyond. Technology for Proteomics and Beyond. *Drug & Market Development*. Pp. 338-343. September 2000.

Russo, E. Proteomics Factories. *The Scientist*. Vol. 14. Pp 1. February 7, 2000.

Valles, C. Glimmer of Hope Emerges in the War against Cancer. Associated Press. Dec. 5, 2000.

Van Brunt, J. Plucking Proteins from Cell Soup. *Signals*. September 28, 2000. (http://www.signalsmag.com)

BIOGRAPHIES

N. Leigh Anderson, Ph.D. is President and CEO of Large Scale Proteomics Corporation, which he co-founded in 1985. His work in the development of 2-D electrophoresis and molecular database technology at Argonne National Laboratory earned him the American Association for Clinical Chemistry's Young Investigator Award for 1982. He holds 10 patents and has written one book and over 100 scientific publications, mainly in the area of proteomics and its applications.

Brian T. Chait, D.Phil. is Camille and Henry Dreyfus Professor and Head of the Laboratory for Mass Spectrometry and Gaseous Ion Chemistry at The Rockefeller University. He directs the NIH-funded National Resource for the Mass Spectrometric Analysis of Biological Macromolecules. Dr. Chait's current research involves rapid identification of proteins, posttranslational modifications of proteins that regulate their function, and the definition of sites of functional interaction between biomolecules.

David Cowburn, Ph.D., D.Sc. is the President and CEO of the New York Structural Biology Center, a new non-profit institute owned by a consortium of New York academic and research centers. His research interests are in the structural biology of signal transduction, using principally nuclear magnetic resonance spectroscopy.

Barbara J. Culliton is Chairman and President of Applera Charitable Foundation. She also serves as Executive Editor of Scientific Communications at Celera Genomics and Executive Editor of Celera's *GenomeNewsNetwork.com,* an international on-line news source for genomics research. Ms. Culliton has an extensive background in scientific publishing and writing including tenures as Deputy Editor of *Nature,* and News Editor for *Science.* She was elected to the Institute of Medicine of the National Academy of Sciences in 1988.

Denis F. Hochstrasser, M.D. is Associate Vice Dean and Professor of Medical Biochemistry and Pharmacology in the Faculty of Medicine at the University of Geneva. He is also President of Clinical Medicine at the university and Head of the Clinical Chemistry Laboratory at Geneva University Hospital. Dr. Hochstrasser oversees all computing activities for the Faculty of Medicine and runs an internationally renowned proteomics research group with a strong focus on bioinformatics. He is a scientific founder of Geneva Bioinformatics SA, and advises a number of major protein-science firms in Europe and the U.S.

Joshua Lederberg, Ph.D. is Sackler Foundation Scholar and President Emeritus of The Rockefeller University. He pioneered in the field of bacterial genetics leading to the Nobel Prize in 1958 and has since become a charter member of the Institute of Medicine of the National Academy of Sciences, and a National Medal of Science recipient. Dr. Lederberg has served on numerous high-level public policy commissions and boards addressing mental health, cancer research, emerging infectious diseases, biological terrorism, space science, and international security.

S. Leslie Misrock, L.L.B has degrees in Chemistry (Columbia University) and Law (Fordham University) and is currently Senior Partner with the law firm of Pennie & Edmonds, L.L.P., specializing in intellectual property issues for inventions in biotechnology and chemistry. Mr. Misrock has extensive litigation experience in this field, including the U.S. Court of Appeals and the International Trade Commission. He is a Fellow of the Royal Chemical Society of London and the American Institute of Chemists.

Rodney W. Nichols has been President and CEO of the New York Academy of Sciences since 1992. Before coming to the Academy he was Executive Vice President of The Rockefeller University and Scholar-in-Residence at the Carnegie Corporation of New York. He has served on the Executive Committee of the Carnegie Commission on Science.

John H. Richards, Ph.D. is Professor of Organic Chemistry and Biochemistry at the California Institute of Technology. His current research aims at a molecular understanding of the mechanisms of protein function in such areas as biological catalysis, the transport of electrons, and the exquisitely specific interactions between proteins and nucleic acids. He also serves as a consultant to a number of biotech companies and research institutes.

B. Michael Silber, Ph.D. is Director of Pharmacogenomics and Clinical Biochemical Measurements at Pfizer Global Research and Development in Groton, CT, and provides leadership for Pfizer's worldwide effort to develop clinical applications and drug discovery in pharmacogenomics, biomarkers, and proteomics.